Sains Malaysiana 54(11)(2025): 2797-2808
http://doi.org/10.17576/jsm-2025-5411-18

An Innovative Approach to Financial Market Analysis: Hybrid ARFIMA with Sieve and Moving Block Bootstrap

(Pendekatan Inovatif untuk Analisis Pasaran Kewangan: Hibrid ARFIMA dengan Saringan dan Butstrap Blok Bergerak)

ALSHAIMAA ELWASIFY[1,2] & ZAIDI ISA[1,2,*]

[1]*School of Mathematical Sciences, Faculty of Sciences & Technology, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia*
[2]*Applied Statistics Department, Faculty of Commerce, Damietta University, 34511 Damietta, Egypt*

Received: 14 April 2025/Accepted: 24 October 2025

ABSTRACT

This paper aims to develop the field of financial time series analysis by focusing on the Egyptian stock market, EGX 30 in particular, using innovative modeling and forecasting techniques. Our study explores the application of ARFIMA models either independently or in combination with advanced bootstrap techniques to improve the accuracy of parameter estimation and forecasting. The study includes four main methodologies: the traditional ARFIMA model, ARFIMA with Sieve Bootstrap (SB), ARFIMA with Moving Block Bootstrap (MBB), and the proposed model that combines the two bootstrap techniques with the ARFIMA model. The proposed model aims to address the time complexities in the financial series, including long term and short-term dependencies. The results show that the proposed model significantly outperforms other traditional and combined models in terms of forecasting accuracy and estimation reliability. This improved performance underscores the importance of integrating advanced bootstrap techniques with traditional models to better understand the complex characteristics of financial data. Our paper contributes to scientific literature by introducing a new approach that has not been applied before in financial markets. It also offers practical applications for investors and financial analysts by providing a robust framework for forecasting and supporting decision-making in dynamic and volatile market environments, with a focus on the Egyptian market. This study represents a basis for applying similar methodologies in other emerging markets.

Keywords: ARFIMA; bootstrap; MBB; Reisen method; sieve bootstrap

ABSTRAK

Kertas ini bertujuan untuk membangunkan bidang analisis siri masa kewangan dengan memberi tumpuan kepada pasaran saham Mesir, khususnya EGX 30, menggunakan teknik pemodelan dan ramalan yang inovatif. Penyelidikan kami meneroka penggunaan model ARFIMA sama ada secara bebas atau digabungkan dengan teknik butstrap lanjutan untuk meningkatkan ketepatan anggaran dan ramalan parameter. Kajian ini merangkumi empat metod utama: model ARFIMA tradisional, ARFIMA dengan Butstrap Saringan (SB), ARFIMA dengan Butstrap Blok Bergerak (MBB) dan model yang dicadangkan yang menggabungkan dua teknik butstrap dengan model ARFIMA. Model yang dicadangkan bertujuan untuk menangani kerumitan masa dalam siri kewangan, termasuk kebergantungan jangka panjang dan jangka pendek. Keputusan menunjukkan bahawa model yang dicadangkan mengatasi model tradisional dan gabungan lain dengan ketara dari segi ketepatan ramalan dan kebolehpercayaan anggaran. Prestasi yang dipertingkatkan ini menggariskan kepentingan mengintegrasikan teknik butstrap lanjutan dengan model tradisional untuk lebih memahami ciri kompleks data kewangan. Kertas kami ini menyumbang kepada kepustakaan saintifik dengan memperkenalkan pendekatan baharu yang belum pernah digunakan sebelum ini dalam pasaran kewangan. Ia juga menawarkan aplikasi praktikal untuk pelabur dan penganalisis kewangan dengan menyediakan rangka kerja yang mantap untuk membuat ramalan dan menyokong proses membuat keputusan dalam persekitaran pasaran yang dinamik dan tidak menentu dengan tumpuan kepada pasaran Mesir. Kajian ini merupakan asas untuk mengaplikasikan metod yang serupa dalam pasaran baharu lain yang muncul.

Kata kunci: ARFIMA; butstrap; butstrap saringan; kaedah Reisen; MBB

INTRODUCTION

Predicting time series data, particularly when considering long-term dependencies and irregular patterns, can provide difficulties. Although many parametric methods have demonstrated their efficacy in numerous instances, they may encounter difficulties when confronted with specific intrinsic properties of time series data. The analysis of time series with long memory characteristics, ARFIMA models are an effective tool for estimating and forecasting these series. However, the most significant limitation of existing methods lies in their difficulty in accurately estimating the parameters, especially in the presence of short-term and long-term dependencies in the data. Parametric methods require a set of prior assumptions about data distribution. The normal distribution is one of these assumptions as it is known, and in practice we find that most researchers ignore the alternative distributions of the normal distribution, which means data miss-specification. Accordingly, there was a need to use parametric inference methods with other non-parametric methods such as bootstrap.

Bootstrap techniques, provide greater flexibility for data analysis without the need for stringent assumptions (Franco, Lana & Reisen 2021). In ARFIMA models, which rely on fractional differentiation, bootstrap techniques are powerful tools for improving the accuracy of estimates and addressing both long- and short-term temporal dependencies, enhancing the reliability of the results (Silva et al. 2006), (Olatayo & Adedotun 2014) (Franco, Lana & Reisen 2021). Examples of such tools include Moving Block Bootstrap (MBB) Sieve Bootstrap (SB) (Lahiri 2003).

This paper makes a good contribution to the current knowledge in the fields of time series analysis and financial forecasting, specifically in the area of modeling the Egyptian stock market EGX30 index. Unlike previous studies that typically combined ARFIMA with only a single bootstrap method, this study distinguishes by its novel integration of the ARFIMA model with multiple bootstrap techniques, which are Sieve Bootstrap and Moving Block Bootstrap together. This hybrid approach enhances the parameter estimation and the accuracy of long-memory forecasting in the presence of complex time dependencies giving stakeholders essential information to make informed decisions in the dynamic and turbulent Egyptian stock market.

In this study, three main techniques were applied: MBB, SB, and the proposed model that combines ARFIMA and both bootstrap techniques, with the aim of improving the parameter estimation and future forecasts of the EGX 30 index. The study highlights the ability of these methods to provide accurate forecasts in dynamic financial environments, which enhances their scientific and practical value.

LITERATURE REVIEW

To provide a comprehensive overview of previous studies, they have been reviewed and classified under 4 main sections, which helps clarify the various trends and comparisons related to the topic.

ARFIMA IN THE ECONOMIC FIELD

Research has shown that these models outperform ARIMA models in dealing with long memory time series. For example, Kurita (2010) analyzed unemployment rate data in Japan and presented a suitable model to represent long memory processes. Paul (2014) confirmed the effectiveness of ARFIMA in forecasting the wholesale price of peas, indicating strong support for the application of this model in business decision making. Omekara, Okereke and Ukaegeu (2016) also found that ARFIMA outperformed ARIMA in forecasting the liquidity ratio of commercial banks in Nigeria, using long memory tests such as GPH. In Safitri et al. (2019), the ARFIMA (1, 1.0572, 3) model was selected to forecast gold prices in Indonesia, highlighting the importance of using techniques such as R/S analysis and GPH approach to estimate fractional difference. Saâdaoui and Rabbouch (2024) presented a hybrid model that combines ARFIMA with LSTM to improve financial time series forecasting, showing greater accuracy and stability compared to conventional models.

ARFIMA IN NON-ECONOMIC FIELDS

ARFIMA models have also been applied to many non-economic fields. For example, Dingari, Reddy and Sumalatha (2019) demonstrated the superiority of ARFIMA using MPL approach in predicting Indian air traffic statistics over conventional ARIMA models. Kartikasari, Yasin and Asih I Maruddani (2020) aimed to estimate COVID-19 pandemic deaths in Indonesia using ARFIMA to address long memory effects, showing high accuracy using ARFIMA (1, 0.431, 0).

BOOTSTRAP TECHNIQUES

For bootstrap techniques, the pioneering work of Efron (1979) is the foundation for these methods, showing that classical resampling produces reliable estimates. However, challenges were observed in the case of temporal data, prompting the development of techniques such as Moving Block Bootstrap (MBB) which has shown great effectiveness in preserving temporal reliability (Vogel & Shallcross 1996). Bühlmann (1997) showed that Sieve Bootstrap outperformed MBB when dealing with long-term temporal dependence, showing lower variance and faster convergence.

BOOTSTRAP TECHNIQUES IN TIME SERIES MODELING

Regarding to time series field, we find Olatayo and Adedotun (2014) used smoothed spectral regression method and truncated geometric bootstrap method for the estimation of the fractional differencing parameter (d). Their results show that the bootstrap is a very good alternative for the estimation of the time series data, especially for the estimation of the fractional parameter of ARFIMA models. Lola, David and Zainuddin (2016) has aimed to improve the performance of the autoregressive (AR) model in forecasting currency exchange rates by developing a hybrid model known as BAR, which combines the AR model with the residual bootstrap technique to reduce the impact of outliers that negatively impact the accuracy of traditional models, where the model has been applied to the 2007 US dollar/Malaysian ringgit exchange rate data, and the results showed that the combination of the two models significantly improved forecast accuracy by reducing the model's sensitivity to outliers, as measured by evaluation indicators such as the root mean square error (RMSE), mean absolute error (MAE), and mean absolute deviation (MAD). This highlights the model's effectiveness in handling volatile financial data. Shang (2023) also compared asymptotic confidence intervals and Sieve Bootstrap for estimating the fractional difference (d), and the results confirmed the higher accuracy of Sieve. Finally, Fokam et al. (2024) proposed a new technique based on Sieve Bootstrap to improve the accuracy of random forests in time series analysis, highlighting the great potential of incorporating bootstrap techniques in improving the forecasting of temporal models.

MATERIALS AND METHODS

This study uses a dataset of 1682 observations to assess the EGX30 index, which represents the Egyptian stock market. To model and forecast the series, ARFIMA models were employed due to their ability to capture long memory behavior. Two bootstrap methods were used to improve parameter estimation and prediction accuracy: Sieve Bootstrap, which handles long-term memory effects, and Moving Block Bootstrap (MBB), which maintains short-term dependency. In order to capitalize on their complementing advantages and enhance the resilience of the ARFIMA model estimates, a hybrid model that combines the two bootstrap approaches was ultimately constructed.

DATA DESCRIPTION

Research focuses on the Egyptian Stock Exchange EGX 30 index. EGX30 is a key Egyptian stock market indicator. The Egyptian market (EGX) manages the EGX30, which tracks the performance of 30 of its most liquid stocks. The EGX30 index is market-capitalization-weighted, therefore, larger companies impact their value. It covers corporations from many industries to show the Egyptian stock market's performance. EGX30 equities are chosen based on market size, liquidity, and trading volume. A comprehensive and varied index that accurately reflects market circumstances is the goal. Starting January 1, 1998, the index was 1000 points. It eliminates bankruptcies, mergers, and acquisitions and includes fully paid ordinary share values. Companies that lose money for three years and violate Egyptian Stock Exchange listing and transparency regulations are excluded (The Egyptian Exchange 2024). Regarding the models used, a detailed explanation is provided.

ARFIMA MODELS

An ARFIMA (p,d,q) process, where p and q are integers and d is a real number, is defined as a stochastic process $\{X_t\}$ described by the equation

$$\phi(B)(1-B)^d X_t = \theta(B)\varepsilon_t \qquad (1)$$

ε_t follows a white noise process WN(0, σ^2). φ(Z) and θ(Z) are polynomials of the back-shift operator B, with degrees p and q, respectively. An ARFIMA (p,d,q) process can be differenced a finite integral number of times until d falls into the interval [-0.5,0.5], at which point it will exhibit stationarity and invertibility (Sun, Chen & Li 2008). When $d = -0.5$, the process is characterized as stationary yet non-invertible. If $-0.5 < d < 0$, the ARFIMA process exhibits intermediate memory or is over-differenced, and the inverse autocorrelations decrease hyperbolically. The ARFIMA process is white noise when d = 0; however, for $0 < d < 0.5$, the ARFIMA (p,d,q) process is stationary with long memory, and the autocovariance function demonstrates hyperbolic decay (Erfani & Samimi 2009).

This study will focus on the smoothed Periodogram (Reisen) approach to estimate the ARFIMA model, which has proven its usefulness in many previous studies such as Reisen (1994) and Silva et al. (2006). This method is well known for its ability to provide accurate parameter estimates, especially in scenarios that require long-memory operations. It also reduces the variance in estimation compared to traditional methods and provides stable and reliable results (Reisen 1994; Reisen, Abraham & Lopes 2001).

SMOOTHED PERIODOGRAM

This approach was initially proposed by Reisen in 1994. A consistent spectrum estimator is employed, utilizing a periodic pattern smoothed with a Parzen lag window to estimate the fractional parameter in the regression equation. Based on the smoothed periodic pattern in Hassler (1993) and Reisen (1994), it is expressed in the following form

$$ln\{f_s(\omega_k)\} = ln\{fu(0)\} - d\,ln\left\{2\,sin\left(\frac{\omega_k}{2}\right)\right\}^2 + ln\left\{\frac{f_s(\omega_k)}{f(\omega_k)}\right\} + ln\left\{\frac{fu(\omega_k)}{fu(0)}\right\} \qquad (2)$$

where

$$f_s(\omega_k) = \frac{1}{2\pi}\sum_{s=-M}^{M} j\left(\frac{s}{M}\right) R(s)\cos s\,(\omega_k) \quad (3)$$

$j\left(\frac{s}{M}\right)$ represents Parzen lag window, an even continuous function in the range $-1 < u < 1$ and $j(0) = 1$, $(u) = j(-u)$ and $M = n^{\beta}$, $0 < \beta < 1$, where $\lim_{n\to\infty, M\to\infty} M/n \to 0$. By determining the range $1 \le k \le g(n)$ and select g(n) as the previous method, the regression equation will be:

$$ln\{f_s(\omega_k)\} \approx ln\{fu(0)\} - d\, ln\left\{2\sin\left(\frac{\omega_k}{2}\right)\right\}^2 + ln\left\{\frac{f_s(\omega_k)}{f(\omega_k)}\right\} \quad (4)$$

where

$$y_k = ln\{f_s(\omega_k)\}, \qquad x_k = ln\left\{2\sin\left(\frac{\omega_k}{2}\right)\right\}^2,$$
$$e_k = n\left\{\frac{f_s(\omega_k)}{f(\omega_k)}\right\}, \qquad a = ln\{fu(0)\}$$

By OLS, the estimation of parameter (d) will be then

$$\hat{d} = -\frac{\sum_{k=1}^{g(n)}(x_k - \bar{x})y_k}{\sum_{k=1}^{g(n)}(x_k - \bar{x})^2} \quad (5)$$

and the variance will be

$$var(\hat{d}) \approx 53928\frac{M}{n\sum_{k=1}^{g(n)}(x_j - \bar{x})^2} \quad (6)$$

The reason for using the Reisen method is that it provides more accurate and reliable estimates compared to other methods such as GPH, especially when the data contains complex features or non-white noise (Reisen 1994; Reisen, Abraham & Lopes 2001).

BOOTSTRAP IN ARFIMA

Bootstrap techniques are more flexible than traditional parametric methods for estimating ARFIMA models, as they do not rely on strict assumptions such as normality or homogeneity, making them more suitable for dealing with long-memory processes. These techniques, such as bootstrapping, help improve the accuracy of estimates and better represent variance through repeated resampling, which enhances the reliability of confidence intervals and inferential statistics (Franco, Lana & Reisen 2021; Paparoditis & Politis 2002).

The lack of assumptions on the underlying data distribution in bootstrap methods offers a distinct advantage when used to ARFIMA models. ARFIMA processes exhibit lengthy memory and persistent dependency, complicating the use of traditional parametric inference methods. The ability to adapt is particularly beneficial when dealing with real-world data that may not meet the ideal circumstances required for parametric methods. Furthermore, non-parametric methods can be tailored to align with the specific attributes of the time series, such as employing the Moving Block Bootstrap or Sieve Bootstrap procedures, which preserve the interdependencies within the data. This tailored methodology enhances the accuracy of conclusions derived from ARFIMA models, providing more reliable estimates of the long memory parameter and other model components.

This study will focus on the use of Sieve Bootstrapping (SB) and Moving Block Bootstrap (MBB) to estimate and forecast data using the ARFIMA model, while improving the accuracy and reliability of the estimates. A proposed model that combines SB and MBB with ARFIMA will also be introduced, leveraging the strengths of each technique and addressing the challenges associated with long and short time dependencies. The model aims to improve the estimations and forecasts, providing an innovative framework for financial time series analysis.

MOVING BLOCK BOOTSTRAP (MBB)

Moving block resampling (MBB) is a technique that divides time series data into overlapping blocks of a predetermined length, and then randomly resamples these blocks with replacement, preserving the temporal structure within each block (Hall 1985; Kunsch 1987). Rather than sampling individual data points, MBB focuses on selecting adjacent blocks of data to create a new series with the same number of observations as the original series. The process is repeated multiple times, allowing the model to be reevaluated using updated parameter distributions (Kreiss & Lahiri 2012). A selector of a simple block length can be obtained from Carlstein (1986).

$$l_{opt} = \left[\frac{6^{.5}\Phi}{1-\Phi^2}\right]^{2/3} n^{1/3} \quad (7)$$

where Φ is estimated of an AR(1) process fitted to the data set.

Sieve Bootstrap

The sieve bootstrap approach for stationary time processes was proposed by Bühlmann (1997) and Kreiss (1992, 1988). Let $X_1,\ldots, X_n$ possess a process mean $E(X_t) = \mu$ and autocovariances $r(k)$, $k \ge 0$, which can be approximated by a stationary autoregressive process $\{Y_t\}$ of autoregressive order p. Subsequently, the equation

$$E\left[(X_t - \mu) - \sum_{j-1}^{p}\beta_j\left(X_{t-j} - \mu\right)\right]^2 \quad (8)$$

can be minimized to obtain the $(\beta_1,\ldots, \beta p)$. Then, the time process $\{Y_t\}$ can be defined as

$$Y_t = \mu + \sum_{j-1}^{p} \beta_j \left(X_{t-j} - \mu\right) + e_t \tag{9}$$

where $\beta \equiv (\beta_1, \ldots, \beta_p)^T = \lceil_p^{-1} r_p$ represents the coefficients of the best-linear predictor of $X_{t-\mu}$ in terms of $(X_{t-1-\mu}, \ldots, X_{t-p-\mu})$, and $\{e_t\}$ are approximately iid random variables with mean $E(e_t) = 0$ and variance $E(e_t^2 = r(0) - \acute{\beta}\lceil\beta$.Further, r_p is a vector of autocovariances defined as $rp = (r(1), \ldots, r(p))^T$and Γp is the $p \times p$ matrix with $r(k-j)$ with the $(k, j)^{th}$ entry of rp. The sieve bootstrap method makes a sample size n bootstrap of $\{Y_t\}$ for implying the distribution of the original time process, $\{X_1, \ldots, X_n\}$. The sieve approximation of $\{Y_t\}$ to $\{X_t\}$ is expected to enhance with an increase in the AR order p relative to the sample size n. The methodology of the sieve bootstrap is delineated as follows (Kim & Kim 2017):

- Define the autoregressive of order $p \equiv pn$ depending on the sample size.
- Estimate the coefficients $(\hat{\beta}_{1n}, \ldots, \hat{\beta}_{pn})$ from observed data $X_1, \ldots, X_n$ using the solution of the sample version of Yule-Walker equations (Brockwell & Davis 1990).
- Define related residuals $\hat{e}_t = (X_t - \bar{X}_n) - \sum_{j-1}^{p} \hat{\beta}_{jn}(X_{t-j} - \bar{X}_n), p+1 \le t \le n.$
- Generate each e_t^* independently from a set of the centered residuals $\{e_t - (n-p)^{-1}\sum_{j=p+1}^{n} e_j : p+1 \le t \le n\}$.
- build the AR-sieve bootstrap replicates from $Y_t^* = \bar{X}_n + \sum_{j-1}^{p} \hat{\beta}_{jn}(Y_{t-j}^* - \bar{X}_n) + e_t^*, t \ge p+1$,where we define $Y_1^* = \ldots = Y_p^* = \bar{X}_n$.

Proposed Model

Combining multiple models is a well-known technique in data analytics and machine learning, where the strengths of each model are leveraged to compensate for the weaknesses of the others (Aladag, Egrioglu & Kadilar 2012). Accordingly, the current model proposes to combine three models so that each improves the performance of the other. The goal of this combination is to achieve better performance than any single model, and to leverage the diversity of models to enhance the results. On this basis, the merging will take place between ARFIMA, Sieve bootstrap (SB), and moving block bootstrap (MBB) to enhance the ARFIMA models.

Step 1: *ARFIMA Model Estimation*

Referring to the equation of ARFIMA model, the ARFIMA (p,d,q) is defined as:

$$\phi(B)(1-B)^d Y_t = \theta(B)\varepsilon_t \quad t = 1,2,\ldots,T \tag{10}$$

B is a back shift operator

$$\phi(B) = 1 - \phi_1 L - \cdots - \phi_p L^p$$

$$\theta(B) = 1 - \phi\theta_1 L - \cdots - \theta_q L^q$$

Estimate ARFIMA parameters: p, d, q any of parametric or semi parametric methods, then get the residuals as follows:

$$Y_t - \hat{Y}_t = \hat{\varepsilon}_t \tag{11}$$

where $\hat{Y}_t$ is the predicted values of ARFIMA model. These residuals contain short-term dependencies that are not fully captured by the ARFIMA model.

Step 2: *Fractional Differencing*

To isolate the long-term dependence from the time series, a fractional differential operator is applied $(1-B)^d$, then

$$Y_t^d = (1-B)^d Y_t \tag{12}$$

Referring to the fractional differencing operator which is defined as

$$(1-B)^d = \sum_{j=0}^{\infty} \binom{d}{j} (-B)^j \tag{13}$$

This process transforms the original time series Y_t into a stationary series Y_t^d, eliminating the effects of long memory.

Step 3: *Sieve Bootstrap (SB)*

The series Y_t^d is stationary now, so the sieve bootstrap can be applied by first fitting an autoregressive model to the series of Y_t^d. After that an AR model is estimated to it as follows

$$Y_t^d = \sum_{i=1}^{p} \Phi_i Y_{t-1}^d + e_t \tag{14}$$

where Φ_i represents the AR coefficient and e_i represents residuals of AR model.

The next step is bootstrapping the residuals e_i extracted from the AR model to generate new samples. Let $e_t^{\cdot}$ the bootstrapped residuals so that

$$e_t^{boot} = resamling\ of\ \{e_1, e_2, \ldots, e_T\} \tag{15}$$

The bootstrapped residuals $e_t^{\cdot}$ and estimated AR coefficients can be used then to fit the new series $Y_t^{d,\cdot}$ which can be defined as

$$Y_t^{d,boot} = \sum_{i=1}^{p} \Phi_i Y_{t-1}^{d,boot} + e_t^{boot} \qquad (16)$$

Step 4: *Applying MBB to residuals*

Referring to the ARFIMA fitted in 1, a moving block bootstrap (MBB) is applied to the residuals capture ϵ_t. This residuals are divided into blocks of size l based on the formula of (Carlstein 1986) in Equation (7), the series of the residuals series will be then

$$\hat{\epsilon}_t = \{\hat{\varepsilon}_1, \hat{\varepsilon}_2, \dots, \hat{\varepsilon}_T\} \qquad (17)$$

And the blocks can be defined as follows

$$B_1 = \{\hat{\varepsilon}_1, \hat{\varepsilon}_2, \dots, \hat{\varepsilon}_l\}, B_1 = \{\hat{\varepsilon}_2, \hat{\varepsilon}_3, \dots, \hat{\varepsilon}_{l+1}\}, \dots \qquad (18)$$

The blocks are randomly resampled with replacement to form a new bootstrapped residual series ϵ_t^{boot}.

Step 5: *Merging the sieve bootstrap series with the MBB-bootstrapped residuals*

In this step, the sieve-bootstrapped differenced series $Y_t^{d,boot}$ from the step 3 is add to the MBB-bootstrapped residuals ϵ_t from step 4. This new series includes both the long-run dependence (from the sieve bootstrap series) and the short-run dependence (from the MBB).

Step 6: *Re-estimation of the ARFIMA parameters from the bootstrapped series*

Finally, using the new bootstrapped series to re-estimate the ARFIMA parameters, then the new ARFIMA model can be defined as

$$\phi(B)(1-B)^d Y_t^{boot} = \theta(B)\varepsilon_t^{boot} \qquad (19)$$

RESULTS AND DISCUSSION

This section deals with the analysis of the results obtained from the application of the different approaches and discussing them to clarify their effectiveness in achieving the objectives of the study.

DATA REPRESENTATION

Herewith is the chart of the EGX 30 (Figure 1), which shows the overall performance of the market during the period from 2013 to the end of 2019 (The Egyptian Exchange 2024), explaining the main trends and changes that occurred in the index during these years.

The chart reflects the performance of the Egyptian Stock Exchange index during the period from 2013 to the end of 2019, as the index witnessed a significant upward trend from 2013 until early 2018, driven by increasing investor confidence and improving economic conditions. The index peaked in 2018, reflecting strong economic activity, followed by a significant decline due to market corrections or negative economic impacts. By 2019, the index was relatively stable with less fluctuations compared to the previous period, indicating a kind of balance in market performance. This performance reflects the dynamism of the Egyptian market and its impact on various economic and political factors during these years.

The plots indicate properties that may fit the ARFIMA model, with the ACF plot showing strong correlations that slowly decay across time periods, reflecting the possible presence of long-term memory in the data. The PACF plot shows a rapid decline after the first lag, indicating the presence of short-term components that can be explained by auto-integrative (AR) processes. These patterns support the application of the ARFIMA model, which combines the properties of long memory and fractional integration to accurately represent and analyze time series (Figure 2).

The first graph (Figure 3) shows the estimation of the fractional integration parameter d under the ARFIMA model across different values of the parameter m, where

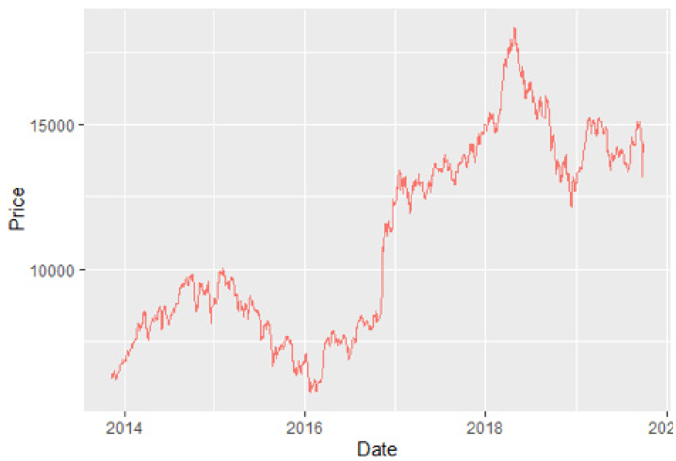

FIGURE 1. Time series of EGX 30

it can be observed that the parameter d stabilizes around a certain value as m increases, supporting the idea of the presence of long-term memory in the time series. The second graph shows the spectral density estimates of the data versus frequency f, with a peak at low frequencies, indicating the effect of long-term memory in the time series. These results support the application of the ARFIMA model for data analysis as it is suitable for representing long-term properties and time-spanning correlations. To ensure the stationarity of the time series, a unit root test was performed and the result can be shown in Table 1.

Table 1 shows the results of the unit root test for the EGX 30 index using ADF, PP, and KPSS tests. The ADF results and Dickey-Fuller value (-1.6811) with p-value (0.7133) showed that the time series is not stationary, as the null hypothesis of unit root was not rejected. Similarly, the PP test supported the result with Dickey-Fuller Z(alpha) value (-61.8141) and p-value (0.7296), indicating that the series is not stationary. On the other hand, the KPSS test showed a level value (1.6919) and p-value (0.01) confirming the rejection of the null hypothesis of stationarity, indicating the presence of unit root. Hence, the time series requires transformations such as differentiation to achieve stability before any further analysis.

ARFIMA MODEL

Several ARFIMA model combinations were tested using the Reisen parameter estimation method to identify the optimal formulation with the best statistical performance. The models were compared according to a set of criteria, including log-likelihood, the AIC and BIC information criteria, and the Ljung–Box test to verify the independence of the residuals. Based on these criteria, ARFIMA (2,d,1) was selected as the optimal model, as it showed the lowest AIC and BIC values and performed well in the Ljung–Box test, indicating its high suitability for representing the time series under study. Table 2 presents the parameter estimates as follows.

All estimates are statistically significant, as evidenced by their low p-values and elevated t-values. These results confirm the importance of these phrases in clarifying the data structure. According to the AIC, BIC, and Ljung-Box test, the ARFIMA (2,0.457,1) model is the optimal selection for modeling both long- and short-term relationships in the dataset.

MOVING BLOCK BOOTSTRAP (MBB)

To implement the MBB, the data is segmented into blocks. The size of the block has been calculated using the Carlstein Equation (7), which is contingent upon the autoregressive parameter of the series. Following the estimation of the autoregressive coefficient AR (1) of EGX 30 at 0.2182 with a p-value of 0.000, it is used in the Carlstein equation to ascertain the optimal block length, which is measured in blocks. After selecting 10 random blocks, the blocks were amalgamated to create a new time series of equivalent length to the original series, and the ARFIMA model was re-estimated utilizing the new series. The procedure was executed 100 times to acquire an adequate quantity of new series, and the ARFIMA model was calculated for each new series. After repeated estimation, we had a set of estimates for each parameter of the model which were used to derive the distribution of the estimated parameters and the upper and lower bounds of the confidence intervals as it is seen in Tables 3.

SIEVE BOOTSTRAP

To implement sieve bootstrap on an ARFIMA model, an autoregressive model was calibrated for the original time series using a fractional difference of $d = 0.457$. The order of the AR model is determined by the characteristics of the residuals and the requisite amount of lags to adequately capture the autocorrelation. Therefore, we implement the AR model on the data with the AR(2) order. The parameters are presented in the Table 4.

Table 4 presents the estimated autoregressive parameters for the original time series, calculated from the fractional difference parameter of the ARFIMA model applied to the original time series. Both lag coefficients are statistically significant. The subsequent stage involves extracting the residuals, which denote the unexplained components of the time series as determined by the model, followed by the application of bootstrap methods. The prediction series will equal the estimated values plus the bootstrapped residuals.

PROPOSED MODEL: HYBRID MODEL

To apply moving block bootstrap to an ARFIMA model, like the sieve bootstrap which an autoregressive model is fitted for the original time series based on the value of the fractional differences d=0.457. Then, we apply the AR model to the time series using the same order AR(2). The parameters can be seen in Table 5. After that, we apply MBB to the residuals resulted from the autoregressive model and t final step to re-estimate the new ARFIMA model with the following parameters.

Table 5 shows the parameters of the hybrid bootstrapped ARFIMA model with the fractional parameter of d = 0.253, indicating moderate long-term memory in the time series. The AR(1) parameter with a coefficient of 0.9979 and a high t-value (58.8) with strong statistical significance indicates a significant influence of one-period prior values on the current values, which reinforces the importance of the prediction based on the first lag. In contrast, the MA(1) parameter has a negative coefficient of -0.7366 and a high negative t-value (-43.8) with strong statistical significance as well, which means that there is a significant influence of random fluctuations in the previous periods on the current values.

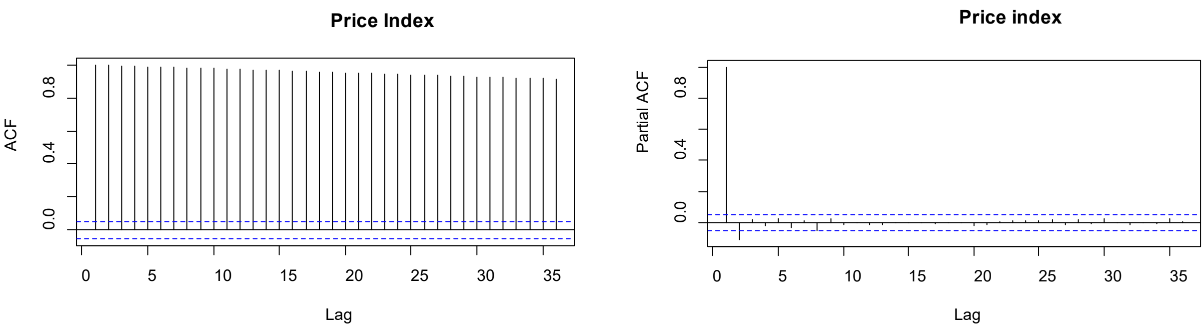

FIGURE 2. ACF and PACF of EGX 30

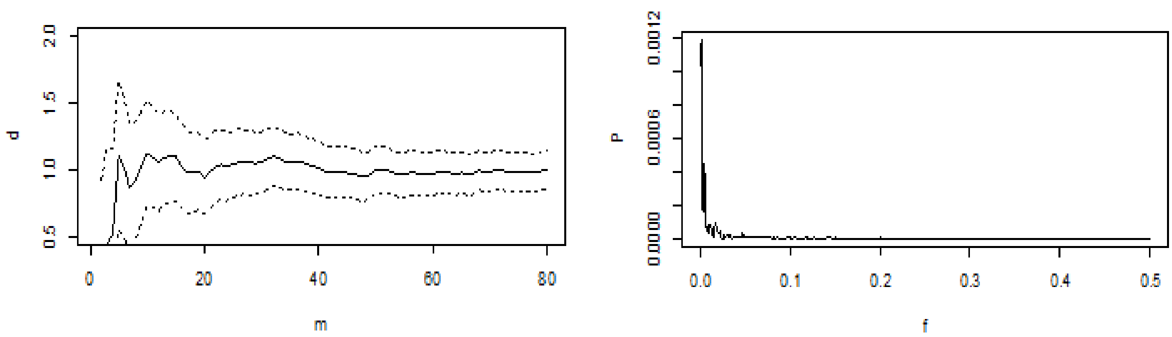

FIGURE 3. Spectral density map and periodogram of EGX 30

TABLE 1. Unit root test of EGX 30

	Dickey-Fuller = -1.6811
ADF Test	Lag order = 9
	p-value = 0.7133
	Dickey-Fuller Z(alpha) = -61.8141
PP Test	Truncation lag parameter = 6
	p-value =0.7296
	KPSS Level = 1.6919
KPSS	Truncation lag parameter = 6
	p-value = 0.01

TABLE 2. The estimated parameters of ARFIMA model on Reisen estimation method

$d = 0.457$				
	Coefficient	S. E	t.value	p.value
Constant	-0.0069	0.0123	-0.56	0.0575
AR (1)	0.7667	0.0708	24.96	0.000
AR (2)	-0.7701	0.0689	-11.18	0.000
MA (1)	-0.8828	0.0558	-15.83	0.000

TABLE 3. The estimated parameters of bootstrapped ARFIMA model

	Coefficient	2.5%	97.5%
d	0.3696	0.6717	0.0675
AR (1)	0.6907	1.1852	0.1961
AR (2)	-0.3628	0.13806	-0.86375
MA (1)	-0.5288	-0.1105	-0.94713

TABLE 4. The estimated parameters of Sieve ARFIMA model

$d-0.457$				
	Coefficient	S. E	t.value	p.value
AR(1)	0.8181	0.0052	15.8	.000
AR(2)	0.1498	0.0052	28.9	.000

TABLE 5. Estimated parameters of Sieve MBB ARFIMA model

$d-0.253$				
	Coefficient	S. E	t.value	p.value
Constant	8.6473	0.0346	25	.000
AR(1)	0.9979	0.0017	58.8	.000
MA(1)	-0.7366	0.0168	-43.8	.000

FORECASTING

The performance of the models was evaluated by testing their out-of-sample forecasting ability for the next 30 days. Forecasting accuracy criteria were used for comparison, as shown in Table 6.

Table 6 presents the evaluation criteria for assessing the forecast accuracy of the EGX 30 series over a 30-day ahead forecast using the standard ARFIMA models, moving block bootstrap (MBB-ARFIMA), sieve bootstrap (SB-ARFIMA) and the proposed model (SB-MBB-ARFIMA). The criteria we have, to compare the accuracy of forecasting at hand are the Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and Mean Absolute Error (MAE). As seen from Table 6, ARFIMA model has the highest error rates: RMSE of 0.0858, MAPE of 0.0027, and MAE of 0.0237, indicating that it performs the least among the models in reducing errors.

The MBB-ARFIMA model improved error rates (RMSE 0.0140, MAPE 0.0013, and MAE 0.0011), demonstrating how bootstrapping enhances ARFIMA by better capturing variance and uncertainty in time series data. Resampling data blocks allows the model to address dependency structures and reduce overfitting, leading to improved forecast accuracy. On the other hand, the Sieve bootstrap ARFIMA model has RMSE of 0.0105, MAPE of 0.0010, and MAE of 0.0084, indicating that the Sieve bootstrap method, which is based on autoregressive structures, improves on ARFIMA by better estimating noise while preserving the fundamental properties of the time series. The best performance is shown with the proposed model across all criteria: RMSE of 0.0102, MAPE of 0.0009, and MAE of 0.0081, where hybrid models typically combine multiple approaches to exploit the individual strengths of each. In this case, the hybrid model leverages ARFIMA with bootstrap techniques to better handle long-term trends and volatility. These results can be supported visually and analytically by plotting the predicted path against the original data, as well as by residual analysis of the proposed hybrid model, as shown in Figure 4.

Figure 4 demonstrates the high degree of accuracy of the proposed model in tracking the dynamic behavior of the ln EGX 30 series, by comparing the estimates (red line) with the actual values (black line) over an out of sample forecast period extending 30 days into the future. The graph shows a near perfect fit between the two series, both in terms of the overall trend and fluctuations, reflecting the model's ability to capture the deep structure of the time series, including trend, cyclical, and volatility components. Besides, the residual analysis showed strong indications of the stability of the proposed hybrid model. The mean

TABLE 6. Forecasting criteria of EGX 30

	RMSE	MAPE	MAE
ARFIMA	0.0858	0.0027	0.0237
MBB – ARFIMA	0.014	0.0013	0.0011
SB – ARFIMA	0.0105	0.0010	0.0084
Hybrid model	0.0102	0.00094	0.0081

FIGURE 4. Actual series of EGX 30 vs. predicted values from the hybrid model

of the residuals was -0.00006, a value very close to zero, indicating the absence of bias in the estimates. Also, the variance of the residuals was 5.34×10^{-5}, a very low value indicating limited error variability.

Based on the analysis of the forecast accuracy indicators, the graph showing the congruence of the forecast path with the original series, and the stability of the residuals, we can conclude that the proposed hybrid model represents the best and most efficient option for representing the dynamics of the time series under study. It not only boasts high forecast accuracy but also stable statistical properties that support its reliability and suitability for practical applications, making it a promising model for analyzing and forecasting long-memory economic variables.

CONCLUSION

In this paper, we have applied five models to time series estimation, including a proposed model based on the combination of two bootstrap techniques: Sieve Bootstrap and Moving Block Bootstrap. The main goal of this work was to improve the parameter estimation of ARFIMA models and increase the accuracy of predictions in the time series. Using the Sieve bootstrap model with ARFIMA is an ideal choice for dealing with time series with long memory. In this model, fractional differentiation is used to remove the long memory from the series, and then an autoregressive (AR) model is estimated for the differentiating series. The residuals of the model are then resampled using the sieve bootstrap, which helps improve the estimates of the d parameter for the long memory. The time series is then reconstructed using the ARFIMA model. One of the main advantages of this model is its ability to handle long-term dependency and increase the accuracy of estimates in ARFIMA models, making it an effective tool for analyzing time series with long memory. However, Sieve bootstrap focuses primarily on long memory, making it less effective for dealing with short-term dependency in the data.

On the other hand, Moving Block Bootstrap is a powerful tool for improving ARFIMA model estimates when there is short-term dependency in the data. The ARFIMA residuals are divided into sequential and overlapping blocks to preserve the time relationships between observations. After the series is divided into blocks, these blocks are resampled with replacement, which allows the short-term time dependency to be preserved. The newly created model is used to improve future predictions and estimate ARFIMA parameters more accurately. This model greatly helps in improving the accuracy of predictions in data with short-term correlations, but it does not handle long-term dependency as well as sieve Bootstrap does.

The proposed model relies on combining two bootstrap techniques to improve the estimation of model parameters. This method takes advantage of the ability of the sieve bootstrap to handle the long-term dependencies present in the time series, while the moving block bootstrap allows preserving the short-term dependencies within the residuals. By combining these two methods, the overall accuracy of the estimates can be improved, and the reliability of the forecasts can be increased. Finally, since the accuracy of forecasts in financial markets is crucial for making investment decisions, the application of this model can provide an effective tool for investors and market analysts, enhancing the chances of making informed decisions and contributing to reducing investment risks based on more reliable forecasts.

REFERENCES

Aladag, C.H., Egrioglu, E. & Kadilar, C. 2012. Improvement in forecasting accuracy using the hybrid model of ARFIMA and feed forward neural network. *American Journal of Intelligent Systems* 2(2): 12-17.

Brockwell, P.J. & Davis, R.A. 1990. *Time Series: Theory and Methods*. 2nd ed. Springer Series in Statistics - Springer.

Bühlmann, P. 1997. Sieve bootstrap for time series. *Bernoulli* 3(2): 123-148.

Carlstein, E. 1986. The use of subseries values for estimating the variance of a general statistic from a stationary sequence. *The Annals of Statistics* 14(3): 1171-1179.

Dingari, M., Reddy, D.M. & Sumalatha, V. 2019. Time series analysis for long memory process of air traffic using ARFIMA. *International Journal of Scientific and Technology Research* 8(10): 395-400.

Efron, B. 1979. Bootstrap methods: Another look at the jackknife. *The Annals of Statistics* 7(1): 1-26.

Erfani, A. & Samimi, A.J. 2009. Long memory forecasting of stock price index using a fractionally differenced ARMA model. *Journal of Applied Sciences Research* 5(10): 1721-1731.

Fokam, C.T., Jentsch, C., Lang, M. & Pauly, M. 2024. AR-sieve bootstrap for the random forest and a simulation-based comparison with rangerts time series prediction. https://doi.org/10.48550/arXiv.2410.00942

Franco, G.C., Lana, G.C. & Reisen, V.A. 2021. Prediction intervals in the ARFIMA model using bootstrap. *Financial Statistical Journal* 4(1): 1-8.

Hall, P. 1985. Resampling a coverage pattern. *Stochastic Processes and their Applications* 20(2): 231-246.

Hassler, U. 1993. Regression of spectral estimators with fractionally integrated time series. *Journal of Time Series Analysis* 14(4): 369-380.

Kartikasari, P., Yasin, H. & Asih I Maruddani, D. 2020. ARFIMA model for short term forecasting of new death cases COVID-19. *E3S Web of Conferences* 202: 13007.

Kim, Y.M. & Kim, Y. 2017. Bootstrap methods for long-memory processes: A review. *Communications for Statistical Applications and Methods* 24(1): 1-13.

Kreiss, J.P. 1992. Bootstrap procedures for AR (∞) - processes. In *Bootstrapping and Related Technique. Lecture Notes in Economics and Mathematical Systems,* edited by Jöckel, K.H., Rothe, G. & Sendler, W. Berlin, Heidelberg: Springer. 376: 107-113.

Kreiss, J.P. 1988. Asymptotic statistical inference for a class of stochastic processes. Verlag nicht ermittelbar.

Kreiss, J.P. & Lahiri, S.N. 2012. Bootstrap methods for time series. *Time Series Analysis: Methods and Applications* 30: 3-26.

Künsch, H. 1987. Statistical aspects of self-similar processes. *Proceedings of the First World Congress of the Bernoulli Society*. Utrecht: VNU Science. 1: 67-74.

Kurita, T. 2010. A forecasting model for Japan's unemployment rate. *Eurasian Journal of Business and Economics* 3(5): 127-134.

Lahiri, S.N. 2003. *Resampling Methods for Dependent Data*. Springer Series in Statistics - Springer.

Lola, M.S., David, A. & Zainuddin, N.H. 2016. Bootstrap approaches to autoregressive model on exchange rates currency. *Open Journal of Statistics* 6(6): 1010-1024.

Olatayo, T.O. & Adedotun, A.F. 2014. On the test and estimation of fractional parameter in ARFIMA model: Bootstrap approach. *Applied Mathematical Sciences* 8(96): 4783-4792.

Omekara, C.O., Okereke, O.E. & Ukaegeu, L.U. 2016. Forecasting liquidity ratio of commercial banks in Nigeria. *Microeconomics and Macroeconomics* 4(1): 28-36.

Paparoditis, E. & Politis, D.N. 2002. The local bootstrap for Markov processes. *Journal of Statistical Planning and Inference* 108(1-2): 301-328.

Paul, R.K. 2014. Forecasting wholesale price of pigeon pea using long memory time-series models. *Agricultural Economics Research Review* 27(2): 167-176.

Reisen, V.A. 1994. Estimation of the fractional difference parameter in the ARIMA (p, d, q) model using the smoothed periodogram. *Journal of Time Series Analysis* 15(3): 335-350.

Reisen, V., Abraham, B. & Lopes, S. 2001. Estimation of parameters in ARFIMA processes: A simulation study. *Communications in Statistics - Simulation and Computation* 30(4): 787-803.

Saâdaoui, F. & Rabbouch, H. 2024. Financial forecasting improvement with LSTM-ARFIMA hybrid models and non-Gaussian distributions. *Technological Forecasting and Social Change* 206: 123539.

Safitri, D., Mustafid, Ispriyanti, D. & Sugito. 2019. Gold price modeling in Indonesia using ARFIMA method. *Journal of Physics: Conference Series* 1217: 012087.

Shang, H.L. 2023. Sieve bootstrapping the memory parameter in long-range dependent stationary functional time series. *AStA Advances in Statistical Analysis* 107(3): 421-441.

Silva, E.M., Franco, G.C., Reisen, V.A. & Cruz, F.R.B. 2006. Local bootstrap approaches for fractional differential parameter estimation in ARFIMA models. *Computational Statistics and Data Analysis* 51(2): 1002-1011.

Sun, R., Chen, Y.Q. & Li, Q. 2008. The modeling and prediction of Great Salt Lake elevation time series based on ARFIMA. *2007 Proceedings of the ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, DETC2007 5 PART B*. pp. 1349-1359.

The Egyptian Exchange. 2024. *EGX30*. https://www.egx.com.eg/en/indexdata.aspx?type=1&nav=1 (Accessed in 2024)

Vogel, R.M. & Shallcross, A.L. 1996. The moving blocks bootstrap versus parametric time series models. *Water Resources Research* 32(6): 1875-1882.

*Corresponding author; email: zaidiisa@ukm.edu.my